

(12) UK Patent Application (19) GB (11) 2 327 548 (13) A

(43) Date of A Publication 27.01.1999

(21) Application No 9715237.5

(22) Date of Filing 18.07.1997

(71) Applicant(s)

British Broadcasting Corporation
(Incorporated in the United Kingdom)
Broadcasting House, LONDON, W1A 1AA,
United Kingdom

(72) Inventor(s)

Peter John Brightwell
Philip Nicholas Tudor
Stephen John Dancer
Oliver Hartwig Werner
Nicholas Dominic Wells

(74) Agent and/or Address for Service

Mathys & Squire
100 Grays Inn Road, LONDON, WC1X 8AL,
United Kingdom

(51) INT CL⁶
H04N 5/268

(52) UK CL (Edition O)
H4F FD3P FD3T FD30K FD30R FGG

(56) Documents Cited
GB 2307613 A

(58) Field of Search
UK CL (Edition O) H4F FGG FGT
INT CL⁶ H04N 5/262 5/265 5/268

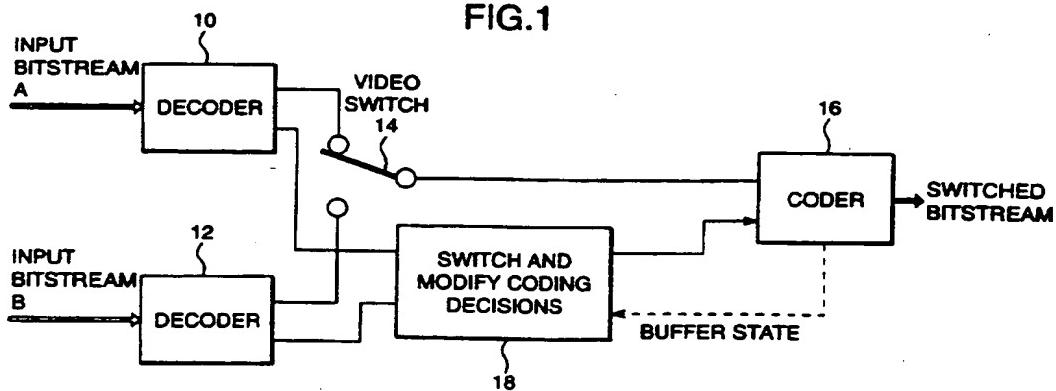
(54) Abstract Title

Switching or splicing compressed video bitstreams

(57) An MPEG bitstream is switched 14 in the video domain in between a decoder and recorder pair which remain continuously in operation but which are transparent through re-use in the recorder of coding decisions taken in the upstream coding of the MPEG bitstream and inferred in the decoder of the switch. Around the switch point, the coding decisions are modified.

The coding decisions supplied to the recorder 16 convey information on some or all of the following parameters: picture dimensions; frame rate; picture structure (frame-coded or field-coded); picture type (I, P or B); whether macroblocks are intra-coded or use prediction; whether forward, backward or bi-directional prediction is used; whether prediction is frame or field based; motion vectors etc.

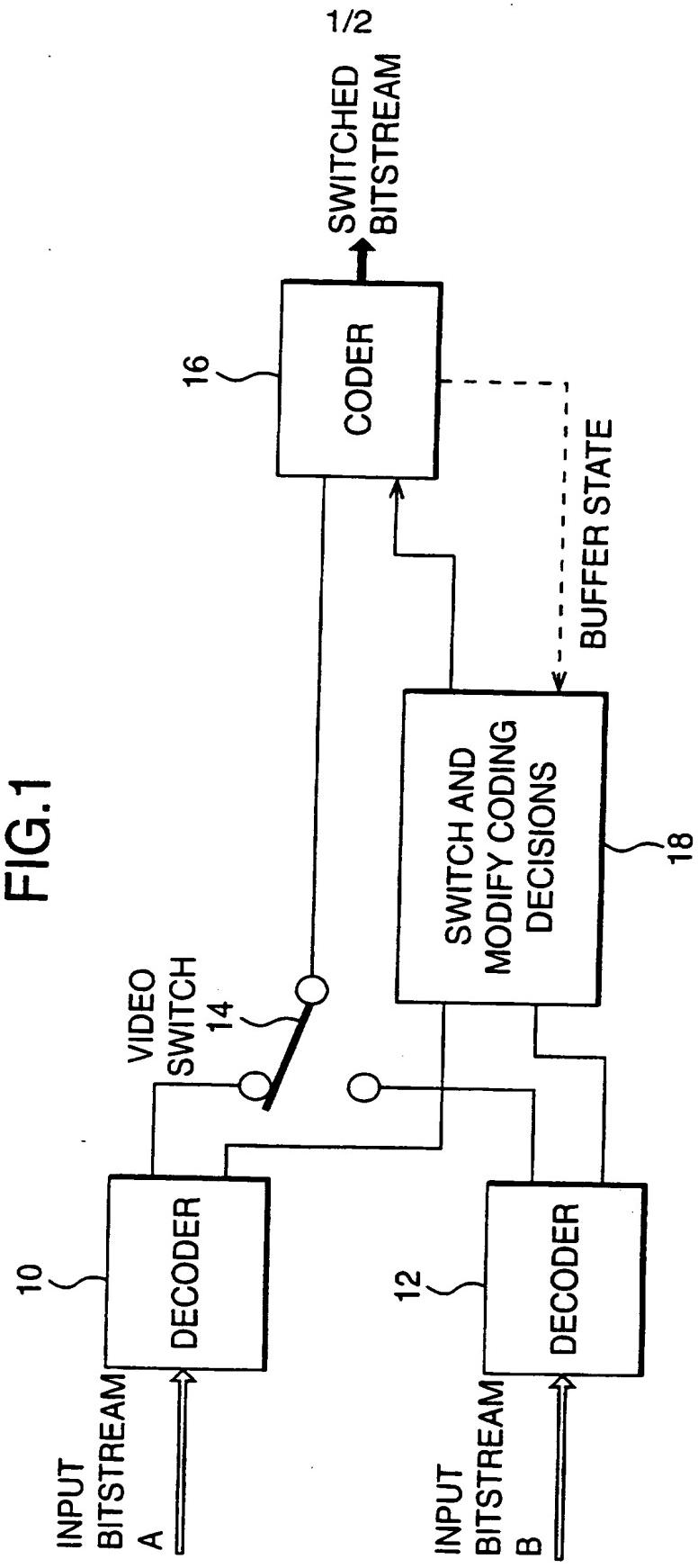
FIG.1



GB 2 327 548 A

At least one drawing originally filed was informal and the print reproduced here is taken from a later filed formal copy.

FIG. 1



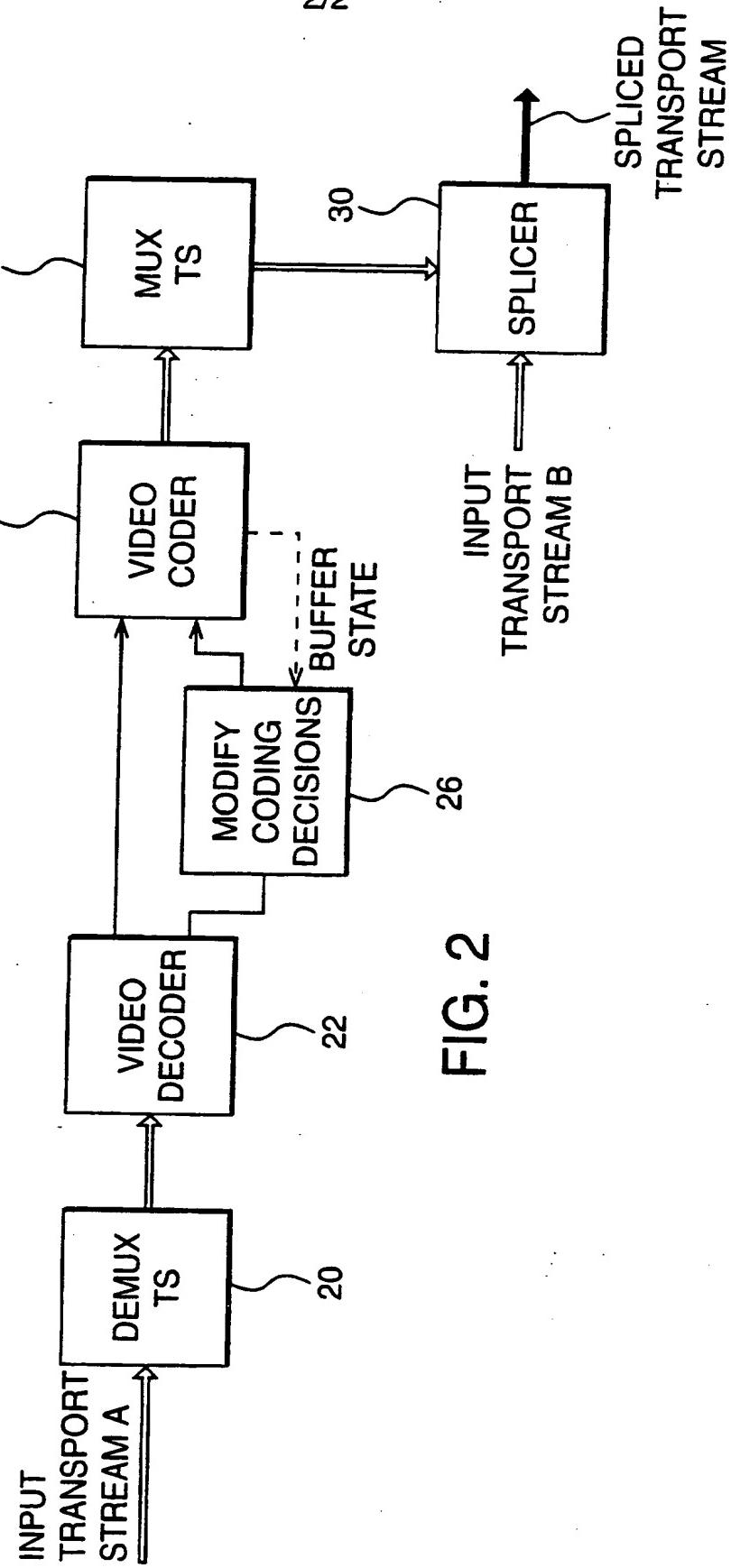


FIG. 2

SWITCHING COMPRESSED VIDEO BITSTREAMS

This invention relates to the field of compressed digital video. As compressed video bitstreams increasingly find their way into the programme chain, techniques for their manipulation are required. One of the most important techniques is the ability to switch between two compressed signals. This is typically required for editing of programmes; this involves switching between different "shots", each of which will be a section from a bitstream. This may be done in real-time or non-real-time. Switching is also required in continuity / presentation, involving real-time switching between different studios or other sources, and for local / regional "opt-out", that is to say switching from a network signal to regional or local programmes or commercials.

Switching of analogue video signals, or of non-compressed digital signals, is relatively straightforward to carry out, as suitable times for switching (called "switching points" here) occur at regular intervals, typically during picture blanking. This is not the case with compressed signals, in which pictures often occupy a variable amount of time and/or bits. Furthermore the compression system may employ temporal prediction, which further complicates switching.

A simple way to switch bitstreams is to decode them, switch in the uncompressed domain, and re-code. This gives good flexibility, but the cascading of coding operations causes loss of picture quality. This can be a serious problem for some types of compression system, where there are many parameters and coding decisions that could be taken differently on re-coding. MPEG-2, currently the most important compression system for broadcasting, is such a system. This invention is intended primarily for MPEG-2 but can be used for other compression systems.

One prior approach to switching without loss of quality is what is known as "transport stream splicing". This involves defining points (called "splicing points") in the input bitstreams (in MPEG transport stream form) at which they can be switched directly. Two variations have been proposed:

"seamless splicing" and "non-seamless splicing".

Splicing is potentially inexpensive to implement, as there is no need to decode or re-code the video. However it lacks flexibility for several reasons.

- The switch can only occur at certain times, determined by the MPEG
5 Group of Pictures (GOP) structure. The frame before the switch must be an I- or P-frame and the frame after it must be an I-frame. This typically means that the switch can only be specified to about half a second of resolution.

- 10 The video switching point determines when any corresponding audio is switched, because the video and audio are part of a single transport stream. It is not possible to have independent video and audio switching points.

The upstream coders must know when switching may be required; if they do not, they might have to insert a large number of splicing points.

- 15 A coder producing a bitstream with seamless splicing points may have to compromise its coding performance to insert these points. This is because the buffer trajectory of a downstream coder must be exactly fixed at the splicing points, making the rate control requirements of the coder more difficult, especially if there are a large number of splicing points. A proposal has been made to get round this problem by inserting synthetic fade-in
20 sequences around the switching points, but this restricts the usefulness of splicing.

Transitions other than simple cuts (ie cross-fades) are not possible.

- Another approach is to switch and re-code the decoded inputs but not make use of the full set of coding options available, for example, within the
25 MPEG-2 "toolset". By reducing the number of decisions and parameters that can change on re-coding, the amount of additional distortion introduced is reduced. One prior proposal makes use of this idea, by employing a relatively simple GOP structure of IBIBIBIB (c.f. IBBPBBPBBPBBIBBP... which is typically used). The drawbacks of this approach are that restricting
30 the coding options tends to require a higher bitrate for the same picture quality, and that it introduces incompatibilities with other coding equipment using the full range of options

An earlier BBC patent application, WO 97/08898 described a bitstream switch using a decode-switch-re-code approach, but in this case, the decoder produced an additional output, giving the coding decisions used by the original coder. By using the same decisions on recoding, the
5 additional degradation introduced would be small. This switch also included bypass paths, used many frames away from the switching point, which ensure that the switch can be made completely transparent, but the use of these complicates the operation of the switch.

10 It is an object of the present invention to provide improved methods and apparatus for the switching of compressed video bitstreams.

Accordingly, the present invention provides in one aspect an apparatus for the switching of compressed video bitstreams comprising a first decoder for receiving a first input bitstream and providing a first decoded video signal and a first coding decision signal; a second decoder for
15 receiving a second input bitstream and providing a second decoded video signal and a second coding decision signal; a video switch for switching between said first and second video signals; a coder for re-coding said switched video signal and coding decision processing means for receiving said first and second coding decision signals and delivering coding decisions to said coder, such that in the steady state said decoding and re-coding is
20 substantially transparent.

25 It has been shown that for re-coding using the same coding decisions, the additional distortion introduced is negligible under normal circumstances. Thus the decoder and re-coder can be kept "in-circuit" at all times, simplifying the switch.

The coding decision information may include: picture dimensions; frame rate; picture structure (frame-coded or field-coded); picture type (I,P or B); whether macroblocks are intra-coded or use prediction; whether forward, backward or bi-directional prediction is used; whether prediction is
30 frame or field based; which field is used; motion vectors; whether field or frame based coding is used; quantiser visibility weighting matrices; quantiser step and buffer state of a downstream decoder.

Advantageously, said coding decision processing means serves around the switching point to modify coding decisions for re-coding. In particular, the refresh strategy may be modified on recoding, by moving the recoding I-frames as P-frames, and/or vice versa, or by use of a "I-P pair" of MPEG
5 frames. The purpose of this is to prevent the downstream decoder's buffer from becoming too full or empty, due to I-frames coming too far apart or too close together in the switched bitstream.

Advantageously, underflow of a downstream decoder buffer is avoided through reduction of bitrate for frames adjacent the switching point. A
10 psycho-visual effect known as "temporal masking" can be employed to prevent the downstream decoder's buffer from becoming too empty. Quantisation noise in the frames very close to the switching point is often not visible to the observer and so the number of bits used in these frames can be deliberately reduced.

15 Preferably, said coding decision processing means receives buffer occupancy information such that the bit usage in the input bitstreams is taken into account in the re-coder's rate control algorithm.

The invention will now be described by way of example with reference to the accompanying drawings, in which:

20 Figure 1 is a block diagram of a switch according to the present invention; and

Figure 2 is a block diagram of apparatus for preconditioning a bitstream for a splice.

25 Figure 1 shows in block diagram form, apparatus for switching between two MPEG-2 video bitstreams. The two decoders each produce two outputs: a decoded video signal, and a coding decision signal. The coding decision signal contains all the relevant information about how the corresponding bitstream was coded that can be deduced from the bitstream. This information includes, but is not limited to, the following: picture dimensions; frame rate; picture structure (frame-coded or field-coded); picture type (I,P or B); whether macroblocks are intra-coded or use prediction; whether forward, backward or bi-directional prediction is used; motion vectors;

quantiser visibility weighting matrices; quantiser step and buffer state of a downstream decoder.

The two decoded video signals are switched as if they were conventional uncompressed signals and the output of the video switch is sent to a special 5 sort of MPEG-2 coder that can make use of a coding decision signal. This coder will reuse some or all of the decisions taken by the coder or coders that created the input bitstreams, as follows.

For frames far away from the switching point, all of the coding decisions are reused. Doing this causes the re-coding process to be near- 10 transparent, i.e. the picture obtained by decoding the output bitstream is virtually indistinguishable from that which would be obtained by decoding the corresponding input bitstream

For frames near the switch point, the MPEG picture type may be changed on re-coding. The main purpose of this is to modify the refresh 15 strategy to be more suitable for the switched bitstream. Typically an intra coded picture (I-picture) is used soon after the switch point, to prevent prediction from occurring "across the cut". More ingeniously, I-pictures in the input bitstreams may be "converted" to non-intra pictures (typically P-pictures), in order to prevent too many I-pictures coming in close 20 succession, which could cause the short-term bit rate to be too high and underflow the buffer of a downstream decoder. Under some circumstances (for instance when there is a large amount of motion in the scene), an intra coded frame may be converted into an intra coded field followed by a forward coded field (to do this the MPEG-2 picture structure is converted 25 from frame to field). This is an alternative method of reducing the short-term bitrate.

The example set out below shows a case where one picture in bitstream A, and two in bitstream B have had their picture type modified on re-coding. These pictures are shown in a bold typeface (for clarity, the pictures are 30 shown in the order in which they are displayed at the decoder output, not the order in which they appear in the bitstream):

	switch point												
Input A	P	B	B	I	I	
Input B	I	B	P	B	B	I	B	B	P
5 Switched output	P	B	B	P	I	B	I	B	B	P	B	B	P

The prediction mode and motion vectors are modified as necessary to take into account any change in picture type, and to prevent predictions from being made "across the cut". Often this involves a simplification; in the above example, the B-picture following the switch point would usually contain bi-directionally predicted macroblocks in input bitstream B, these are modified to be forward predicted, and the backward vectors discarded. However, where an I-picture is converted to a P-picture, new vectors will need to be estimated, unless MPEG-2 concealment vectors were available in the input bitstream.

The quantiser parameters are modified in order to control the number of bits produced on re-coding. This is typically done by the rate controller in any MPEG-2 coder, which monitors the buffer state of a downstream decoder (actually it monitors the state of its own output buffer, which mirrors that of the decoder buffer).

The rate controller for this invention differs from a conventional one in that the bit usage of the input bitstreams is conveyed via the coding decisions, and is used to set targets for the number of bits to be produced on re-coding. These targets are then used to determine the quantiser parameters. Also, over a number of frames following the switch point, the rate controller attempts to bring the buffer state of a downstream decoder to match what it would have been if bitstream B had been sent directly to the decoder. Typically the number of bits required is set to be slightly lower than the input bitstream. When the decoder buffer state matches, the rate controller enters a "locked" mode in which the quantiser parameters (as well as all other coding decisions) are kept exactly the same as in the input

bitstream. After this point, the switch is virtually transparent. Due to an effect called "temporal masking", the level of noise in the pictures immediately before and after the switch point can often be allowed to be higher than at other times. The eye is distracted by the change in picture material and fails to notice the additional noise. Thus the number of bits used in these pictures may deliberately be made quite small, if this helps in making the buffer state match as above.

In addition, the effect of psycho-visual masking is greater in certain areas of the picture, for example around the edges of the picture, than in other areas, for example in the centre, or near brighter objects. Thus, bitrate may be further reduced by permitting quantisation noise to rise more in areas of the picture where the effect of psycho-visual masking is greater than in other areas.

Bitrate may be reduced for several frames in the vicinity of the switch point, for example about 1-10 frames either side of the switch point, and not necessarily the same number of frames on either side.

In certain cases, it may be necessary to add bits to the bitstream to maintain a desired decoder buffer occupancy; this can be achieved by padding with dummy information, or by inserting portions of data to be transmitted with the picture (if any) at this point in the bitstream.

Although this specific description has focused on the switching of MPEG-2 bitstreams, many of the techniques are applicable for use with other types of DCT-based compressed video bitstreams, such as JPEG, ETSI and MPEG-1, or even non-DCT-based bitstreams, such as in wavelet or fractal-based systems.

The invention can easily be modified to cope with the case when switching between compressed and non-compressed inputs; here the coder makes use of the coding decisions in the compressed input, but makes its own decisions when the non-compressed input is selected

The switch can be extended to incorporate a reduction in bitrate, in other words, the rate of the output bitstream is lower than one or both of the input bitstreams. For example, a lower bitrate may be used for transmission to the

viewer than is used for distribution from the studio to the transmitter. In such cases, the switch is not intended to be transparent, and there is no need to attempt to match the downstream decoder's buffer state as above.

5 In addition to performing simple switches, other video transitions may be required. The most important of these is the cross-fade. MPEG coding of cross-fading sequences tends to give poor quality pictures; this is due to the difficulty in estimating motion vectors and poor performance of forward prediction. This invention can be extended to perform cross-fades.

10 For this purpose, the described video switch takes the form of a video cross-fading device. It will be recognised that during the fade, decisions from both input bitstreams are available for use in re-coding.

15 In one coding strategy, the decisions from bitstream A are used for the first part of the fade, and for the last part, the decisions from bitstream B are used. In an alternative strategy, the potential coding quality is assessed separately using each set of decisions, and the better set chosen.

Temporal masking is not appropriate, as there is no sudden change from one sequence to the other. However, the mixing of the sequences tends to decrease visibility of quantisation noise, and the rate controller of the coder can exploit this fact.

20 Some techniques of this invention can be used to precondition a transport stream as part of a splicing device.

25 Referring now to Figure 2, this shows splicing from transport stream A to transport stream B. Transport stream A is demultiplexed to elementary bitstreams (only the video bitstream is shown). The video bitstream is decoded to video with coding decisions. When many frames from a splice, the video is re-coded with the same decisions and remultiplexed; this will be transparent under normal conditions. Near the time of splicing, the number of bits produced will be adjusted to meet the buffer constraints for splicing, typically by adjusting the quantiser parameters.

30 In addition to the above, the inputs must be synchronised to ensure that splicing occurs on the correct picture type, and modifications are required to the time stamps and clock reference information in the transport stream.

These aspects are not shown in the above diagram.

Another possible embodiment of the present invention is a non-real-time edit conforming switch based in software. To speed up the operation of this, some modifications may be made to the essentially hardware approach described thus far. Bitstreams may be directly copied when the switch is effectively transparent. Only a partial decode – as far as the DCT domain – may be performed under some circumstances, provided that the picture type is not changed. This may lead to a small loss but acceptable loss in quality.

CLAIMS

1. Apparatus for the switching of compressed video bitstreams comprising a first decoder for receiving a first input bitstream and providing a first decoded video signal and a first coding decision signal; a second decoder for receiving a second input bitstream and providing a second decoded video signal and a second coding decision signal; a video switch unit for switching between said first and second video signals; a coder for re-coding said switched video signal and coding decision processing means for receiving said first and second coding decision signals and delivering coding decisions to said coder, such that in the steady state said decoding and re-coding is substantially transparent.
5
2. Apparatus according to Claim 1, wherein said coding decision signals convey information on some or all of the following parameters: picture dimensions; frame rate; picture structure (frame-coded or field-coded); picture type (I,P or B); whether macroblocks are intra-coded or use prediction; whether forward, backward or bi-directional prediction is used; whether prediction is frame or field based; which field is used; motion vectors; whether field or frame based coding is used; quantiser visibility weighting matrices; quantiser step and buffer state of a downstream decoder.
15
3. Apparatus according to Claim 1 or Claim 2, wherein said coding decision processing means serves around the switching point to modify coding decisions for re-coding.
20
4. Apparatus according to Claim 3, wherein said coding decision processing means serves around the switching point to re-code I-frames as P-frames, and vice versa.
25

5. Apparatus according to Claim 3, wherein said coding decision processing means serves around the switching point to convert an intra coded frame into an intra coded field followed by a forward coded field.
5. Apparatus according to any one of the preceding claims, wherein underflow of a downstream decoder buffer is avoided through reduction of bitrate for one or more frames in the vicinity of the switching point.
7. Apparatus according to Claim 6, wherein bitrate is reduced by permitting quantisation noise to rise to a level just masked by the psycho-visual effect of a marked change in picture content.
10. 8. Apparatus according to Claim 6 or 7, wherein bitrate is reduced by permitting quantisation noise to rise more in areas of the picture where the effect of psycho-visual masking is greater than in other areas.
15. 9. Apparatus according to any one of the preceding claims, wherein said coding decision processing means receives buffer occupancy information such that the bit usage in the input bitstreams is taken into account in the re-coder's rate control algorithm.
20. 10. Apparatus according to any one of the preceding claims, wherein the coder for re-coding said switched video signal passes into an unlocked rate control mode on switching and re-enters a locked mode, in which coding decisions are used directly, once the determined buffer state of a downstream decoder matches that applicable to a steady state of the switched video signal.
11. Apparatus according to any one of the preceding claims, wherein said video switch unit is adapted to perform cross fades.

12. Apparatus according to Claim 11, wherein the unit is arranged to assess the potential coding quality using decisions based on each input bitstream and to select the set of decisions giving better coding quality.

13. Apparatus for preconditioning a bitstream for a splice comprising a decoder for receiving an input bitstream and providing a decoded video signal and a coding decision signal; a coder for re-coding said video signal and coding decision processing means for receiving said coding decision signal and delivering coding decisions to said coder, wherein said coding decision processing means serves around the intended splice point to modify coding decisions for re-coding.

14. Apparatus according to Claim 13, wherein said coding decision signals convey information on some or all of the following parameters: picture dimensions; frame rate; picture structure (frame-coded or field-coded); picture type (I, P or B); whether macroblocks are intra-coded or use prediction; whether forward, backward or bi-directional prediction is used; whether prediction is frame or field based; which field is used; motion vectors; whether field or frame based coding is used; quantiser visibility weighting matrices; quantiser step and buffer state of a downstream decoder.

15. Apparatus according to Claim 13 or Claim 14, wherein said coding decision processing means serves around the splice point to re-code I-frames as P-frames, and vice versa.

16. Apparatus according to Claim 12 or Claim 13, wherein said coding decision processing means serves around the splice point to convert an intra coded frame into an intra coded field followed by a forward coded field.

17. Apparatus according to any one of Claims 13 to 16, wherein underflow of a downstream decoder buffer is avoided through reduction of bitrate for one or more frames in the vicinity of the splice point.
- 5 18. Apparatus according to Claim 17, wherein bitrate is reduced by permitting quantisation noise to rise to a level just masked by the psycho-visual effect of a marked change in picture content.
- 10 19. Method for the switching of compressed video bitstreams, comprising the steps of decoding a first input bitstream and providing a first decoded signal and a first coding decision signal; decoding a second input bitstream and providing a second decoded signal and a second coding decision signal; switching between said first and second video signals; and re-coding said switched video signal utilising information from said first and second coding decision signals.
- 15 20. Method according to Claim 19, wherein said input bitstreams are decoded to DCT coefficient level.
- 20 21. Method according to Claim 19, wherein said input bitstreams are decoded to video signals.
- 25 22. Method according to Claim 21, wherein said coding decision signals convey information on some or all of the following parameters: picture dimensions; frame rate; picture structure (frame-coded or field-coded); picture type (I,P or B); whether macroblocks are intra-coded or use prediction; whether forward, backward or bi-directional prediction is used; whether prediction is frame or field based; which field is used; motion vectors; whether field or frame based coding is used; quantiser visibility weighting matrices; quantiser step and buffer state of a downstream decoder.

23. Method according to Claim 21 or Claim 22, wherein said coding decisions are modified around the switching point.
24. Method according to Claim 23, wherein said coding decision processing means serves around the switching point to re-code I-frames as P-frames, and vice versa.
- 5 25. Method according to Claim 23, wherein said coding decision processing means serves around the switching point to convert an intra coded frame into an intra coded field followed by a forward coded field.
- 10 26. Method according to any one of Claims 19 to 25, wherein underflow of a downstream decoder buffer is avoided through reduction of bitrate for one or more frames in the vicinity of the switching point.
27. Method according to Claim 26, wherein bitrate is reduced by permitting quantisation noise to rise to a level just masked by the psycho-visual effect of a marked change in picture content .
- 15 28. Method according to Claim 25 or Claim 27, wherein bitrate is reduced by permitting quantisation noise to rise more in areas of the picture where the effect of psycho-visual masking is greater than in other areas.
29. Apparatus substantially as herein described, with reference to the accompanying drawings.
- 20 30. A method substantially as herein described, with reference to the accompanying drawings.



Application No: GB 9715237.5
Claims searched: 1-30

Examiner: John Coules
Date of search: 5 September 1997

Patents Act 1977
Search Report under Section 17

Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

UK Cl (Ed.O): H4F FGG,FGT

Int Cl (Ed.6): H04N 5/268,5/262,5/265

Other:

Documents considered to be relevant:

Category	Identity of document and relevant passage	Relevant to claims
X	GB 2307613 A (BBC) see 54,56,58 and 64 in fig 4, and page 18 line 25 to page 19 line 9.	1,13 and 19 at least

X	Document indicating lack of novelty or inventive step	A Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E Patent document published on or after, but with priority date earlier than, the filing date of this application.